This lecture covers induction, which is the most common technique for proving universally quantified sentences over the natural numbers and other discrete sets of objects (lists, trees, tilings, graphs, programs, etc.). We need induction, or something like it, because any attempt at proof by enumeration of possible worlds is doomed to failure—with infinitely many objects, there are infinitely many possible worlds that can be defined.

The principle of induction

The principle of induction is an axiom of the natural numbers. Recall from Lecture 1 that the natural numbers satisfy Peano's axioms (where we have written $n+1$ instead of $s(n)$ for clarity):

> 0 is a natural number
> If n is a natural number, $n+1$ is a natural number

These axioms don't quite *define* the natural numbers, because they include no statement of the kind, "... and by the way, there are no other natural numbers." This makes it impossible to use these axioms alone to prove that any property holds for all natural numbers n.

The principle of induction essentially fills this gap. Informally, it says the following:

> If you can prove that some property holds for 0, and you can prove that the property holds for $n+1$ if it holds for n, then you have proved that the property holds for all the natural numbers.

To state this principle formally, let $P(n)$ be an arbitrary proposition about the natural number n. (For example, $P(n)$ might be "$2^n > n$".) The principle of induction is as follows:

Axiom 3.1 (Induction): For any property P, if $P(0)$ and $\forall n \in \mathbf{N}\ (P(n) \implies P(n+1))$, then $\forall n \in \mathbf{N}\ P(n)$.

This says that if $P(0)$ holds, and $P(0) \implies P(1)$, and $P(1) \implies P(2)$, and so on, then $P(n)$ must be true for all n. This seems pretty reasonable. (Later we will see that the axiom has an alternative form that seems even more reasonable.)

Inductive proofs

The standard first example is to verify the well-known formula for the sum of the first n positive integers:

Theorem 3.1: $\forall n \in N \ \sum_{i=1}^{n} i = n(n+1)/2$

[Note: $\sum_{i=1}^{n} i$ means $1+2+\cdots+n$ but is more precise. When $n=0$ the summation has no terms so is 0; when $n=1$ it has just the term $i=1$ so the sum is 1.]

In many, but not all, cases of proof by induction, the property P used in the induction is exactly the one we want to prove. That is the case here: $P(n)$ is the proposition that $\sum_{i=1}^{n} i = n(n+1)/2$.

BASE CASE
INDUCTIVE STEP
INDUCTIVE HYPOTHESIS

To construct an inductive proof, first we need to establish the **base case** $P(0)$, then we need to prove that $P(n+1)$ follows from $P(n)$. The latter proof is called the **inductive step**, and usually involves all the work. The key idea of induction, though, is that this step is easier than proving the whole universal proposition from scratch because you get to *assume* that $P(n)$ is true when proving $P(n+1)$. This assumption is called the **inductive hypothesis**.

Proof: The proof is by induction over the natural numbers.

- Base case: prove $P(0)$.
  $P(0)$ is the proposition $\sum_{i=1}^{0} i = 0(0+1)/2$. By the definition of summation, this is equivalent to $0=0$, which is true.

- Inductive step: prove $P(n) \implies P(n+1)$ for all $n \in \mathbf{N}$.

  1. The inductive hypothesis is $\displaystyle\sum_{i=1}^{n} i = n(n+1)/2$.
  2. To prove: $\displaystyle\sum_{i=1}^{n+1} i = (n+1)(n+2)/2$.
  3. Re-expressing the summation for $n+1$ in terms of the summation for n (*for which we have the formula*), we have

$$\begin{aligned}\sum_{i=1}^{n+1} i &= \left(\sum_{i=1}^{n} i\right) + (n+1) \text{ by the definition of summation} \\ &= n(n+1)/2 + (n+1) \text{ using the inductive hypothesis} \\ &= n(n+1)/2 + 2(n+1)/2 = (n+1)(n+2)/2\end{aligned}$$

Hence, by the induction principle, $\forall n \in \mathbf{N} \ \displaystyle\sum_{i=1}^{n} i = n(n+1)/2$. $\square$

This example shows the format of an inductive proof, which you should follow religiously. It also exhibits the technique common to all inductive proofs: the proposition $P(n+1)$ is reformulated into a part that comes directly from $P(n)$, for which we already have the answer, and an "extra bit" that is then combined with the answer from $P(n)$ to give the answer for $P(n+1)$.

Example proofs

Theorem 3.2: *$\forall n \in \mathbf{N}$, $n^3 - n$ is divisible by 3.*

It will be helpful to define "divisible by" more precisely. We write $a|b$ (a divides b, or b is divisible by a); the definition is

Definition 3.1 (Divisibility): For all integers a and b, $a|b$ if and only if for some integer q, $b = aq$.

As in the previous example, we take the definition of P straight from the theorem; $P(n)$ is the proposition that $3|(n^3 - n)$.

Proof: The proof is by induction over the natural numbers.

- Base case: prove $P(0)$.
  $P(0)$ is the proposition that $3|(0^3 - 0)$ or $3|0$, which is true from the definition of divisibility with $q = 0$.

- Inductive step: prove $P(n) \implies P(n+1)$ for all $n \in \mathbf{N}$.

  1. The inductive hypothesis is $3|(n^3-n)$, or $n^3-n=3q$ for some integer q.
  2. To prove: $3|((n+1)^3-(n+1))$. We do this by showing that $(n+1)^3-(n+1)=3r$ for some integer r.
  3. Re-expressing the quantity for $n+1$ in terms of the quantity for n (*for which we know a simple formula*), we have

$$\begin{aligned}(n+1)^3-(n+1) &= n^3+3n^2+3n+1-(n+1) \text{ expanding out the cubic term}\\ &= (n^3-n)+3n^2+3n \text{ rearranging to isolate the known quantity}\\ &= 3q+3(n^2+n)=3(q+n^2+n)\end{aligned}$$

  4. Hence, since q and n are integers, we have $3|((n+1)^3-(n+1))$.

Hence, by the induction principle, $\forall n \in \mathbf{N}\ 3|(n^3-n)$. □

Again, the trick is to reduce $P(n+1)$ to $P(n)$ and a little extra fact; here the little extra fact is that $3|(3n^2+3n)$. Sometimes, this step can be carried out by simple manipulation to pull $P(n)$ out of $P(n+1)$; sometimes, you really have to understand what's going on at a deep level.

The next example proves an inequality between two functions of n; such inequalities are useful in computer science when showing that one algorithm is more efficient than another. Notice that the base case is $P(2)$ rather than $P(0)$. This is an obvious variant[1] of the standard principle of induction; we can begin with any fixed integer to show that all subsequent integers satisfy some property. (If we want to prove $P(n)$ for *all* integers, including negative ones, we must do two inductions—one upwards from, say, 0 and one downwards from 0.)

Theorem 3.3: $\forall n \in N\ \ n>1 \implies n! < n^n$

Proof: The proof is by induction over the natural numbers greater than 1.

- Base case: prove $P(2)$.
  $P(2)$ is the proposition that $2! < 2^2$, or $2<4$, which is true.

- Inductive step: prove $P(n) \implies P(n+1)$ for all natural numbers $n>1$.

  1. The inductive hypothesis is $n! < n^n$.
  2. To prove: $(n+1)! < (n+1)^{(n+1)}$
  3. Re-expressing the quantity for $n+1$ in terms of the quantity for n (*for which we know a simple inequality*), we have

$$\begin{aligned}(n+1)! &= (n+1)\cdot n! \text{ by the definition of factorial}\\ &< (n+1)\cdot n^n \text{ using the inductive hypothesis}\\ &\qquad\qquad \text{and } \forall x,y,a\ \ a>0 \wedge x<y \implies ax<ay\\ &< (n+1)\cdot(n+1)^n \text{ using } \forall x,y,a\ \ x>0 \wedge a>0 \wedge x<y \implies x^a<y^a\\ &= (n+1)^{(n+1)}\end{aligned}$$

Hence, by the induction principle, $\forall n \in \mathbf{N}\ n>1 \implies n! < n^n$. □

[1] There is a way to avoid this new "variant." We can let $P(n)$ be $n>1 \implies n! < n^n$, and prove the base case $P(0)$ trivially because the premise $0>1$ is false. When we prove the inductive step, of course, we will need to prove $P(0) \implies P(1)$ and $P(1) \implies P(2)$ simply by proving $P(1)$ and $P(2)$ separately, and then prove $P(n) \implies P(n+1)$ for every $n>1$. The work done is exactly the same but the proof looks messier.

Not a proof

Theorem 3.4: *All iMacs are the same colour.*

Proof: We rephrase as: $\forall n \in \mathbf{N}$ if $n > 0$ then every set of n iMacs is monochromatic. The proof is by induction over the number of iMacs in a set.

- Base case: prove $P(1)$.
  $P(1)$ is the proposition that every set of 1 iMacs is monochromatic. This is obviously true.
- Inductive step: prove $P(n) \implies P(n+1)$ for all natural numbers $n > 0$.
  1. The inductive hypothesis states that every set of n iMacs is monochromatic.
  2. To prove: every set of $n+1$ iMacs is monochromatic.
  3. For each set $S = \{i_1, \ldots, i_{n+1}\}$ of $n+1$ iMacs, consider the n-element subsets $S_1 = \{i_1, \ldots, i_n\}$ and $S_2 = \{i_2, \ldots, i_{n+1}\}$. These sets have $n-1$ elements in common and their union is S.
  4. By the induction hypothesis, S_1 is monochromatic. Hence, i_1 is the same colour as i_n.
  5. By the induction hypothesis, S_2 is monochromatic. Hence, i_n is the same colour as i_{n+1}.
  6. Hence all elements of the set S are the same colour, i.e., every set of $n+1$ iMacs is monochromatic.

Hence, by the induction principle, all iMacs are the same colour. □

Absurd as the conclusion may seem, this proof is not an easy one to diagnose because the inductive step *is* correct for a "typical" case where, say, $n = 3$. It really is true that if every set of 3 elements is monochromatic, then every set of 4 elements is monochromatic; and so on for all larger n. We must look to the root. We know that $P(1)$ is true but $P(2)$ is false, so the error must come in the inductive step when $n = 1$. In that case, S has two elements and the n-element subsets are the singletons $\{i_1\}$ and $\{i_2\}$. Step 4 is correct since i_1 and i_n are the same object. The error comes in step 5: although we say (correctly) that $S_2 = \{i_2, \ldots, i_{n+1}\}$, this does *not* mean that i_n is an element of S_2! More intuitively, the inductive step relies on S_1 and S_2 having a nonempty intersection so that colour can be "propagated" from S_1 to S_2. When S has 2 elements, this is not the case.

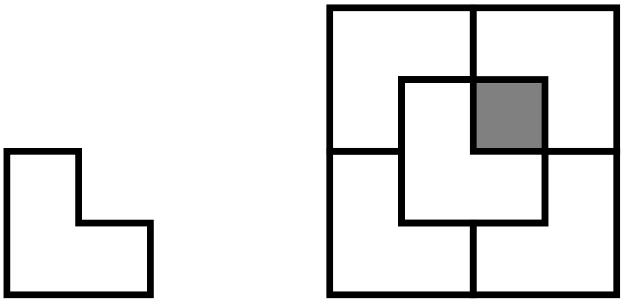
Strengthening the inductive hypothesis

In all the previous examples, we have taken the property $P(n)$ directly from the statement of the theorem to be proved. This does not always work; sometimes the inductive hypothesis fails to provide sufficient "juice" to connect one property to the next. Let's consider an example.

TILINGS

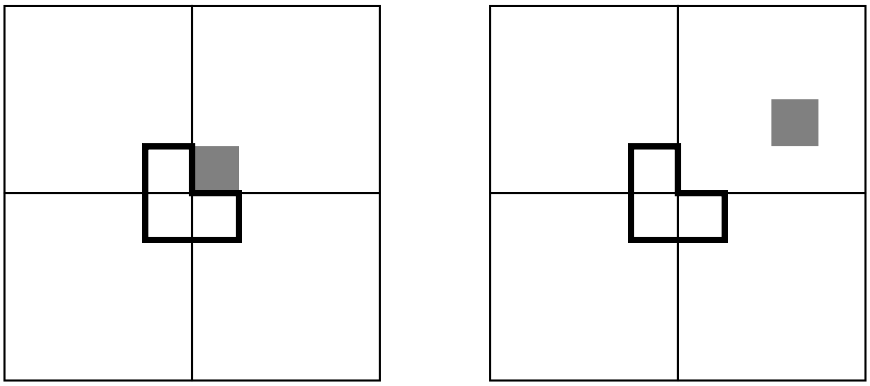
The study of **tilings** is a mathematical pastime with serious implications for circuit design, solid-state physics, and other areas. A tiling is a complete and exact covering of a region (usually planar) with multiple copies of some smaller shape. One of the best-known results is that an 8×8 region with two opposite corner squares removed *cannot* be tiled by 2×1 dominoes.

In our case, a binary-obsessed computer science department wants to build its new building around a courtyard of size $2^n \times 2^n$ for some n. An inept state contractor has supplied the department with L-shaped tiles to pave the courtyard (Figure 1). An alert student in CS 70 proves that $\forall n\ 3 | 2^{2n} - 1$, so any tiling is going to leave one hole. "No matter," says the ever-alert department chair, "We'll leave one hole at the center for a flattering statue of the yet-to-be-found building donor," little knowing that this would be Pokemon Toys Inc.



(a) (b)

Figure 1: (a) An L-shaped tile. (b) An L-tiling of a 4×4 region leaving one hole in a center square.



(a) (b)

Figure 2: (a) The inductive step for the first proof attempt: a region of size $2^{n+1} \times 2^{n+1}$ with one central hole can be divided into four regions of size $2^n \times 2^n$, whose four central corners can be covered with one hole and one L-tile. (b) The inductive step for the second proof attempt: a region of size $2^{n+1} \times 2^{n+1}$ with one hole anywhere can be divided into four regions of size $2^n \times 2^n$; one of the four regions contains the hole, and the three central corners of the other regions can be covered with one L-tile.

Theorem 3.5: *$\forall n \in \mathbf{N}$, every region of size $2^n \times 2^n$ with one central hole has an exact L-tiling.*

Proof: (Attempt 1) The proof is by induction over the natural numbers.

- Base case: prove $P(0)$.
  $P(0)$ is the proposition that a 1×1 region with one central hole has an exact L-tiling. This is true, requiring 0 tiles.
- Inductive step: prove $P(n) \implies P(n+1)$ for all $n \in \mathbf{N}$.
  1. The inductive hypothesis states that every region of size $2^n \times 2^n$ with one central hole has an exact L-tiling.
  2. To prove: every region of size $2^{n+1} \times 2^{n+1}$ with one central hole has an exact L-tiling.
  3. The $2^{n+1} \times 2^{n+1}$ region can be divided into four regions of size $2^n \times 2^n$, whose four central corners can be covered with one hole and one L-tile (see Figure 2(a)).
  4. Each of the four remaining regions is of size $2^n \times 2^n$ with one hole. Therefore, by the inductive hypothesis, it has an exact L-tiling.

5. If a set of disjoint regions has an exact tiling, their union has an exact tiling. (Note: this requires a separate induction!) Therefore, every region of size $2^{n+1} \times 2^{n+1}$ with one central hole has an exact L-tiling.

Oops. Our use of the inductive hypothesis in step 4 is flawed! The inductive hypothesis guarantees an L-tiling for regions with a *central* hole, not a *corner* hole.

□

When an impasse of this kind occurs, you should consider *strengthening* the theorem you are trying to prove. This is rather counterintuitive—if you fail to prove a theorem, can it make sense to prove a harder one??

Proving a stronger theorem has the advantage of strengthening the inductive hypothesis $P(n)$, which may enable a proof for $P(n+1)$. But notice that it also strengthens $P(n+1)$, which can be a disadvantage. Sometimes it works out, sometimes it doesn't. As yet, there is no recipe for success beyond "Insight!" or "Practice!"

For this example, the obvious strengthening is to drop the restriction to *central* holes:

Theorem 3.6: *Every region of size* $2^n \times 2^n$ *with one hole has an exact L-tiling.*

This is a stronger theorem because the hole is now allowed to be anywhere rather than just in the middle; the old theorem is a special case. If we knew this fact $P'(n)$ for a particular n, we would certainly be able to complete the inductive step for the old $P(n+1)$ in the above proof attempt, because now we can tile regions with corners missing. This illustrates the advantage of strengthening the theorem.

Unfortunately, we are no longer interested in proving the old $P(n+1)$—"Every region of size $2^{n+1} \times 2^{n+1}$ with one central hole has an exact L-tiling." Instead, we must prove the stronger $P'(n+1)$—"Every region of size $2^{n+1} \times 2^{n+1}$ with one hole (*anywhere*) has an exact L-tiling." Fortunately, we can do this given $P'(n)$. The proof is illustrated in Figure 2(b); its detailed completion is left as an exercise.

There are at least two more interesting aspects of this example. First, the inductive proof translates very naturally into a recursive algorithm for *constructing* an L-tiling. Second, unlike previous proofs in these notes, the proof is actually *informal*. The structure of the inductive argument is perfectly OK, but the detailed claims are supported by appeal to a picture rather than to general axioms. This is because we have not supplied any axioms for two-dimensional shapes, tilings, dividing regions, etc. For simple cases, we can rely on human intuition on these topics, but when it comes to tiling 11-dimensional regions with 9-dimensional holes using 10-dimensional tiles (as in string theory or robotic path-planning), one *must* fall back on an axiomatic basis.